Ashton, D., & Trinh, C. R. (2018). Evaluating the information content of earnings forecasts. *Accounting and Business Research*, *48*(6), 674-699. https://doi.org/10.1080/00014788.2017.1415800

Peer reviewed version

Link to published version (if available):
10.1080/00014788.2017.1415800

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Taylor & Francis at
https://www.tandfonline.com/doi/full/10.1080/00014788.2017.1415800?scroll=top&needAccess=true. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/

Evaluating the Information Content of Earnings Forecasts

David Ashton
Department of Accounting and Finance
School of Economics, Finance and Management
University of Bristol

Chau (Ruby) Trinh[*†]
Department of Accounting and Finance
School of Economics, Finance and Management
University of Bristol

October 2017

Abstract: This study develops a framework to compare the ability of alternative earnings forecast approaches to capture the market expectation of future earnings. Given prior evidence of analysts' systematic optimistic bias, we decompose earnings surprises into analysts' earnings surprises and adjustments based on alternative forecasting models. An equal market response to these two components indicates that the associated earnings forecast is a sufficient estimate of the market expectation of future earnings. To apply our framework, we examine four recent regression-based earnings forecasting models, alongside a simple earnings-based random walk model and analysts' forecasts. Using the earnings forecasts of the model that satisfies our sufficiency condition, we identify a set of stocks for which the market is unduly pessimistic about future earnings. The investment strategy of buying and holding these stocks generates statistically significant abnormal returns. We offer an explanation as to why this and similar strategies might be successful.

Keywords: Earnings forecasts, earnings response coefficient, market expectation of future earnings, portfolio selection, analysts' forecasts.

JEL Codes: G11, G17, M40, M41

[*] Corresponding author. *E-mail address: chau.trinh@bristol.ac.uk*
[†] We thank Edward Lee, Mark Clatworthy, Nikola Petrovic, Sheila Ellwood, Colin Clubb, Daniela Acker, Christian Leuz, Beatriz García Osma, William Rees, two anonymous reviewers and participants at the 2015 European Accounting Association Doctorial Colloquium for helpful comments and suggestions. All remaining errors are our own.

1. Introduction

Investors use a variety of sources of information to form their expectations of a firm's future earnings and stock returns. Undoubtedly one of the more influential sources is analysts' reports. Analysts have access to a wider variety of information than most investors, including personal contacts with senior executives of companies. Nonetheless, or perhaps because of these contacts, their one-year-ahead forecasts are known to have a systematic optimistic bias. Although regression-based models use a much more restricted information set, their proponents argue that their more objective and rigorous treatment of information is likely to lead to a more reliable forecast of earnings. There is no consensus on what is the best proxy for the market expectation of future earnings.

Our study focuses on providing a theoretical framework for identifying superior forecasting ability in a set of contrasting models. We apply our framework to six different forecasting models, including those made by analysts. The one-year-ahead forecasts of one of the models are found to be more informative than analysts' forecasts about the market expectation of future earnings. We then exploit this superiority to implement an investment strategy that is shown to consistently generate significant abnormal returns.

Knowing the market expectation of future earnings is important for use in portfolio formation, implied cost of capital estimation and investors' behaviour studies. However, owing to the unobservable nature of the market expectation, a generally accepted method is to study the earnings response coefficient (ERC), which is the price reaction to errors in the earnings forecast (Ball and Brown 1968). Nevertheless, much of the work in this area has proceeded on an ad-hoc empirical basis using basic regression techniques without a theoretical foundation.[1] Although we also use a linear regression-based approach, we develop a theoretical framework by which the relative contribution of corrections to analysts' forecast errors can be evaluated in terms of the correlation structure of errors in forecasts. Using this framework, we explore the relationship of unexpected returns to two measures of earnings surprise: analysts' forecast errors and adjustments based on alternative forecasting models.[2] The equal earnings response coefficient of these two measures is evidence that the model-based forecast alone provides a sufficient estimate of the market expectation of future earnings.

To evaluate our new approach, we use four recent regression-based earnings forecasting models alongside a simple earnings-based random walk (RW) model and analysts' forecasts. The regression-based models are (1) the Hou, van Dijk and Zang (HDZ) model (Hou et al. 2012); (2) the SO model (So 2013); (3) the Konchitchki, Lou, Sadka and Sadka (KLSS) model (Konchitchki, et al. 2013); and (4) the Harris and Wang (HW) model (Harris and Wang 2013). To predict future earnings, these models use different combinations of subsets of publicly available information, ranging from accounting information and market information to analysts' forecasts.[3] All the regression-based models are claimed to outperform analysts' forecasts in terms of forecast accuracy and/or the ERC. Our analysis is primarily concerned with the relative merits of alternative one-year-ahead earnings forecast proxies since financial statements are announced on a yearly basis and accounting information-based models can therefore provide a forecast once a year at most.

[1] Ohlson (1991) comments: 'Another theoretical problem concerns the relevance of unexpected earnings as a variable explaining returns. This construct appears to have the status of a "folklore concept" with limited economic content'.

[2] The former error equals the deviation of analysts' forecast from the actual reported earnings while the latter is the deviation of model-based forecasts from analysts' forecast.

[3] The HDZ and SO models base their earnings forecasts purely on accounting information. The KLSS model uses both market information (past stock returns) and analysts' forecasts. The HW model is adapted by Harris and Wang (2013) using the theoretical approach in Ashton and Wang (2013). This model uses both accounting and market information to predict future earnings.

We decompose earnings surprise proxies associated with these earnings forecasts into analysts' forecast errors and adjustments based on alternative forecasting models. Regressing unexpected returns on these two components shows that only the KLSS model has an equal ERC for these two components, satisfying the condition laid out in our framework. It implies that the KLSS model is the superior proxy for the market expectation of one-year-ahead earnings. Further investigation reveals that this KLSS model is effectively an optimal combination of the two benchmarks – analysts' forecasts of earnings and a RW model – against which regression-based forecasts have normally been separately evaluated.

We examine the economic evidence for our findings by applying them to a strategy of portfolio formation. While analysts' forecasts are the most forward looking, the deviations of the KLSS forecasts, a historical data-based model, from those of analysts reflect the expected market's correction for analysts' errors. A significant negative deviation might signal over-adjustments and an unduly pessimistic forecast of one-year-ahead earnings. Thus, if this is true, buying and holding these stocks should generate statistically significant abnormal returns. Hence, we implement a buy-and-hold policy for the stocks in the bottom 10% of the differences between earnings forecasts of the KLSS model and the consensus forecasts of analysts. The results show that this strategy, on average, generates statistically significant positive abnormal returns of 3.4% for one-month buy-and-hold investments and 8.8% for one-year buy-and-hold investments.

In summary, our paper contributes to the literature by developing a rigorous formal framework for comparing earnings forecasts of financial analysts and other models in terms of capturing the market expectation of future earnings. By understanding the statistical implication of observed values of ERC in terms of the correlation between forecast errors, we are able to identify a forecasting model, the KLSS model, which contains more information about one-year-ahead forecasts of market expectations than that of analysts. We show that this forecast is effectively based on an adjusted combination of benchmarks against which regression-based forecasts have hitherto been measured. We exploit the information content of the KLSS forecast, which is additional to that of analysts, to produce a viable investment strategy. Finally, we offer a new explanation as to why one-year-ahead regression-based forecasting models may offer profitable investment strategies.

The remainder of the paper is organized as follows. Section 2 presents the theoretical analysis that guides our empirical investigations. Section 3 describes the model structure and data. Section 4 contains the empirical results, while section 5 describes an investment strategy based on these results. Section 6 concludes and outlines the implications of our study.

2. Motivation and theoretical foundations for the value of information

2.1. Motivation

Earnings data are an important input for fundamental valuation and asset-allocation decisions (Hou et al. 2012). Accurate earnings forecasts help to improve the quality of these decisions while better estimates of the market expectations of accounting earnings are important for many reasons. First, they can be used in the context of investment analysis. If one is able to predict future reported earnings precisely and shows that these forecasts are different from the market's expectation of future earnings embedded in price, it should be possible to generate excess returns. Second, the market expectations of future earnings are used in the estimation of implied cost of capital (e.g. Claus and Thomas 2001, Gebhardt et al. 2001, Easton and Monahan 2005), which is an important input for capital appraisal investment and financial budgeting decisions. Any estimation errors of the market earnings expectation will lead to inaccurate estimation of the implied cost of capital and suboptimal investment decisions.

While the forecast quality can be assessed based on forecast bias or accuracy (Foster 1977, Fama and French 2000, 2006, Hou et al. 2012, So 2013), determining whether such a forecast contains valuable information about the market expectation of future earnings is more challenging because the market expectation is unobservable. A more generally accepted method is to examine the price reaction to errors in the earnings forecast (ERC). The implication is that the higher the ERC, the better the model is at capturing the market expectation of earnings (Beaver 1970, Fried and Givoly 1982, Brown et al. 1987a, O'Brien 1988, Hou et al. 2012).

Based on the ERC measure, early researchers found analysts' forecasts outperformed the random walk and univariate time-series models (Brown and Rozeff 1978, Fried and Givoly 1982, Brown et al. 1987a, b). Hence, they have been extensively used as a proxy for the market expectation of future earnings in the context of the estimation of earnings surprise (Brown et al. 1987a, Walther 1997) and the implied cost of capital (Claus and Thomas 2001, Gebhardt et al. 2001, Easton et al. 2002, Botosan and Plumlee 2005, Easton and Monahan 2005). However, more recent literature has argued that analysts' forecasts are less accurate than the simple RW model under some conditions (Bradshaw et al. 2012, Gerakos and Gramacy 2013) and regression-based models (such as Hou et al. 2012, So 2013). Analysts forecasts also contain both systematic forecast bias and prediction errors (Abarbanell and Bushee 1997, Frankel and Lee 1998, Bradshaw et al. 2001, Easton and Sommers 2007, Hughes et al. 2008, Dichev and Tang 2009, Bradshaw 2011, Bosquet et al. 2015).

Given the possibly predictable nature of analysts' forecast errors, investors might rationally seek to adjust analysts' forecasts for their biases (Feldman et al. 2003). Hughes et al. (2008) and So (2013) estimate the predictable component of analysts' errors as the difference between more accurate earnings forecasts and analysts' forecasts. They argue that if the strategy of sorting firms by the predicted errors fails to generate abnormal returns, investors know analysts' errors. Hughes et al. (2008) find that the market is able to predict analysts' forecast errors, while So (2013) concludes that the market fails to do so and that the market still overweighs analysts' forecasts. Other researchers (Hou et al. 2012, Harris and Wang 2013, Li and Mohanram 2014) claim that their regression-based forecasts outperform analysts' forecasts in terms of capturing the market expectation of future earnings.

Explanations for these controversial findings remain limited due to the lack of a solid theoretical foundation for model evaluation. We therefore develop a new approach to evaluate whether a model is a sufficient estimate of the market earnings expectation by exploring the relationship between the unanticipated earnings surprises based on analysts' forecasts, and any adjustment that can be provided by additional information contained in regression-based forecasts. This is the subject of the discussion in the next section.

2.2. Information measures and earnings response coefficients

The random variables E_{t+1}^{rep}, E_t^m, E_t^a and E_t^f represent the reported earnings at time $t+1$ for an individual firm, the market expectation of earnings at time t, the consensus analysts' forecast at time t and a forecast made by a regression model, respectively.[4] We note that realizations of E_{t+1}^{rep}, E_t^a and E_t^f are observable but that the market expectation of earnings E_t^m is not strictly observable.[5] In the case where the market expectation of earnings (E_t^m) is different from the actual reported earnings (e_{t+1}^{rep}), we are likely to see a price reaction in response to the earnings surprises (ES_{t+1}), as shown in equation (1):[6]

$$ES_{t+1} = e_{t+1}^{rep} - E_t^m \quad . \tag{1}$$

[4] Because all empirical work concerns the aggregate behaviour of the variable across firms, at this stage firm subscripts are omitted.

[5] E_t^m is not strictly observable yet could be implicitly studied through the stock price reaction during earnings announcements.

[6] We adopt the usual mathematical convention that capitals denote random variables and the corresponding lower-case letters the realization of those variables.

We assume that E_t^m is a noisy proxy for e_{t+1}^{rep} and that the market expectation prior to the actual announcement of earnings is normally distributed about actual reported earnings. Similarly, forecasting models produce noisy information about future reported earnings. We can model these processes, as in equation (2):[7]

$$\begin{aligned} E^{rep} - E^m &= -(\varepsilon^m) \\ E^{rep} - E^a &= -(\eta + \varepsilon^a) \\ E^{rep} - E^f &= -\varepsilon^f \\ E^a - E^f &= \eta + \varepsilon^a - \varepsilon^f \end{aligned}, \tag{2}$$

where ε^m, ε^f and ε^a denote error terms assumed normal with zero means and variances of σ_m^2, σ_f^2 and σ_a^2, respectively, and where $\rho_{ma}, \rho_{mf}, \rho_{af}$ denote the correlations between the respective error terms, while $\eta \geq 0$ represents a positive bias or optimism on the part of analysts. The assumption of zero means and normality of the model-based forecast errors justifies the regression process and will be evidenced in the empirical work that follows. The assumption of normality gives us a linear form from the relationship between the expected surprise $(E^{rep} - E^m)$ and observed values of e^{rep}, e^a and e^f:[8]

$$\mathrm{E}\{E^{rep} - E^m \mid e^{rep},\ e^a,\ e^f\} = (\lambda_1 - \lambda_2)\eta + \lambda_1(e^{rep} - e^a) + \lambda_2(e^a - e^f), \tag{3}$$

where λ_1 and λ_2 are coefficients associated with analysts' forecast errors and expected corrections for analysts' forecast errors. Theoretical values of these coefficients are derived in Appendix A.

We further assume that the price reaction during the earnings announcements, or unexpected return (*UR*), follows a mean variance model, with the expected one-year-ahead price response being subject to an uncertainty adjustment:

[7] For reasons of clarity, at this stage time subscripts are omitted.

[8] The unitalicized, unsubscripted E followed by {} in equations (3) and (4) denotes the expectations operator.

$$UR = k_1 \times \mathrm{E}\{E^{rep} - E^m \mid e^{rep},\ e^a,\ e^f\} = \gamma_0 + \gamma_1 \times (e^{rep} - e^a) + \gamma_2 \times (e^a - e^f), \quad (4)$$

where k_1 is the pricing multiple, $\gamma_1 = k_1 \times \lambda_1$ and $\gamma_2 = k_1 \times \lambda_2$.

We expect this intercept (γ_0), which is a combination of a measure analysts' optimism η in equation (3) and the price attached to uncertainty, to be positive. In this form, γ_2 measures the relative information of the correction to analysts' forecasts contained in the difference $e^a - e^f$. Equation (4) can also be written as:

$$UR = \gamma_0 + (\gamma_1 - \gamma_2)(e^{rep} - e^a) + \gamma_2(e^{rep} - e^f). \quad (5)$$

Here the ratio $(\gamma_1 - \gamma_2):\gamma_2$ denotes the relative contributions of analysts' forecasts to the regression-based forecasts in predicting the market earnings surprise. Our theoretical modelling in Appendix A offers insights into the regression structure in terms of more familiar measures of forecasting performance. It also helps us to interpret the results of our empirical investigations when we use regression equations (4) and (5). The most important of these is the relationship between γ_1 and γ_2, which are just pricing multiples of λ_1 and λ_2 in Appendix A:

$$\gamma_1 = k_1 \times \frac{(\rho_{ma} + \rho_{mf})}{1 + \rho_{af}} \sigma_m \quad (6)$$

$$\gamma_2 = k_1 \times \frac{\sigma_m(\rho_{mf} - \rho_{ma}\rho_{af})}{1 - \rho_{af}^2} \quad (7)$$

$$\gamma_1 - \gamma_2 = k_1 \times \frac{(\rho_{ma} - \rho_{mf}\rho_{af})\sigma_m}{1 - \rho_{af}^2}. \quad (8)$$

If analysts' forecasts are superior to most regression-based forecasts as measured by their correlation coefficients, then $\rho_{ma} \geq \rho_{mf} > 0$ and $0 < \rho_{af} \leq 1$. This implies $\rho_{ma} - \rho_{mf}\rho_{af} \geq 0$, which in turn leads to $\gamma_1 \geq \gamma_2 > 0$.

A positive value of γ_2 implies that the regression-based model potentially contains useful forecasting information in addition to analysts' one-year-ahead forecasts about likely future price reactions. The relative weight, $(\gamma_1 - \gamma_2):\gamma_2$, or relative importance of the analysts' forecast generated by equation (5), is a decreasing function of the accuracy of the regression forecast relative to that of analysts, where the accuracy of both is measured by their respective correlation coefficients (ρ_{ma}, ρ_{mf}) with the market.

As proved in Appendix A, $\gamma_1 = \gamma_2$ implies that the regression-based forecast by itself is a sufficient statistic for analysts' forecasts. In this case, the relative weight given to the analysts' forecast in equation (5) is zero and the price reaction can be explained purely in terms of the error in the regression-based forecast.

3. Model structure and data selection

3.1. Model structure

At time t, in order to predict earnings at time $t+1$, we assume that all our investors who rely on regression-based models study the relation between earnings and its determinants from a restricted information set ($\mathbf{IS}$), as in equation (9):

$$e_t^{rep} = \alpha + \boldsymbol{\beta}.\mathbf{IS}_{t-1} + \varepsilon_t, \tag{9}$$

where e_t^{rep} represents reported earnings at time t, $\mathbf{IS}_{t-1}$ is the information vector at time $t-1$, α is the intercept and ε_t is the residual term. We use the estimated coefficients $\left\{\hat{\alpha}, \hat{\boldsymbol{\beta}}\right\}$ applied to the information set $\mathbf{IS}_t$ to make a forecast e_t^f of next period's earnings using the information available at time *t*:

$$e_t^f = \hat{\alpha} + \hat{\boldsymbol{\beta}}.\mathbf{IS}_t. \tag{10}$$

Earnings surprises (forecast errors) associated with one-year-ahead forecasts (e_t^f) are then estimated when actual reported earnings (e_{t+1}^{rep}) are announced, as in equation (11):

$$es_{t+1} = e_{t+1}^{rep} - e_t^f \ . \tag{11}$$

We select four recent studies which use different combinations of the information set, including accounting information, market information and analysts' forecasts, as detailed in Appendix B. Interestingly, three out of the four models claim to be better than analysts in terms of forecast accuracy and/or earnings response coefficients. This is despite the fact that they use publicly available accounting and finance information, which appears to be just a subset of what analysts use (Brown et al. 2015). We also investigate a random walk model without drift. As a benchmark against which the performance of the regression-based models is assessed, we use analysts' earnings per share (EPS) forecasts from I/B/E/S. A comparison of analysts' forecasts with the others should provide evidence of the degree to which different information sets predict market expectations of the future earnings. It also potentially provides evidence or otherwise of analysts' competence, who are considered to be dominant information intermediaries in capital markets (Frankel et al. 2006).

3.2. Data selection

The sample includes all NYSE, AMEX, and NASDAQ listed securities with a fiscal year-end of 31st December and sharecode 10/11 at the intersection of the CRSP monthly returns file, the COMPUSTAT fundamental annual file and the I/B/E/S consensus EPS forecast files from January 1983 to December 2015. We use Cusip8 to merge the two databases. The start date is chosen due to the low availability of analysts' forecasts prior to 1983 and the sample period is similar to that of Hou et al. (2012) and Konchitchki et al. (2013). To mitigate the effect of outliers, earnings and other variables each year are winsorized at the 1st and 99th percentiles. We return to this issue later in this section.

We follow Konchitchki et al. (2013) and So (2013) in using only companies with December fiscal year-ends. Analysts' forecasts are collected in April each year, allowing for a reporting lag of three months. This ensures that neither regression-based models nor analysts have an information advantage over the others, since information about these companies as well as analysts' forecasts is announced at approximately the same time. Compared with that of Hou et al. (2012), our research design provides a better match between the return window and the horizon of the expected earnings measure, as well as aligning analysts' forecast accuracy.[9] Furthermore, as well documented, varying seasonal and investor sentiment affects stock price and excess returns are positively correlated with such shifts in sentiment (Rozeff and Kinney 1976, Lee et al. 2002). Hence, using only December year-end firms, whose earnings are announced during a similar period, allows us to isolate the impact of economic conditions and/or investor sentiment on the return-earnings regression. Because of this fiscal year-end requirement, our sample accounts for approximately 55% of the whole population. As a result, generalizability could be limited because December year-end firms are typically larger than non-December year-end firms and associated with different levels of price reaction (Smith and Pourciau 1988, Bamber et al. 2000).

Table 1 shows the number of firm-year observations for each model at every stage, including the sample used, earnings forecasts, earnings forecasts after merging with other models, scaling and matching with annualized abnormal stock returns. The regression-based models, HDZ and HW, unsurprisingly have greater coverage than analysts. The KLSS model has the poorest coverage since it requires a minimum of two consecutive years of actual EPS

[9] Hou et al. (2012) generate forecasts at end of June for all firms with different fiscal year-ends and compare these with the latest analysts' forecast. Hence, for firms with a fiscal year-end of July, their forecasts and analysts' forecast are just a month away from the actual earnings announcement, yet they are then matched with annualized abnormal returns for model valuations. Their results are subject to this mismatch between the return window and the horizon of the earnings surprise measure. In our study, forecasts of one-year-ahead earnings are generated in April, so are approximately a year from the next annual earnings announcements. Earnings surprises associated with these forecasts are then matched to annualized abnormal returns at the model-valuation stage.

data together with the corresponding analysts' forecast. We obtain a sample of 27,903 firm-year observations, which contains observations that are available across models and abnormal stock returns. Nevertheless, while analysts aim to forecast the actual I/B/E/S earnings, the statistical comparison with regression-based modellers is based on winsorized values, placing an unfair disadvantage on analysts. Hence, the sample is later reduced to 26,506 when observations with winzorized earnings are removed to provide a level playing field for analysts' and model-based earnings forecasts. The need for a common dataset probably results in a concentration of larger companies with established analyst coverage.

<Insert Table 1 about here>

Table 2 presents summary time-series average statistics of the variables used in the cross-sectional regression earnings models for the period 1983–2015. Note the differences in the units of measurement. Panel A describes statistics of the HDZ model where earnings, total assets, dividends and accruals are unscaled factors. Compared with the values in Hou et al. (2012), the means of earnings and total assets are higher, probably due to the shorter sample period. The summary statistics in Panel B are different from those in Harris and Wang (2013) due to more censoring of the data, omitting observations with extreme accounting values, removing firms with negative book values and penny stocks. The summary statistics associated with the KLSS model in Panel C are similar to those reported in Konchitchki et al. (2013). Panel D presents statistics of variables used in the SO model.[10]

<Insert Table 2 about here>

3.3. Cross-sectional earnings regression

---

[10] Hou et al. 2012 (who use all firms with a sample period from 1968 to 2008) have a mean income before extraordinary items (IB) of $49.07m and mean total assets (AT) of $1529.78m. Konchitchi et al. 2013, who use December-fiscal-year-end firms with a similar sample period (1985–2010), have a mean change in EPS deflated by stock price (CIB) of 0.003, a mean lagged one-year return (RET) of 0.194, and a mean analysts' forecast of changes in EPS deflated by stock price (CAF) of 0.019. Harris and Wang (2013) (1963–2011) report a mean earnings per share (IBPS) of $0.685, a mean book value per share (B) of $9.105, and a mean adjusted stock price (APRC) of $13.51. So (2012) does not present summary statistics.

To establish the relationship between earnings at time t and the information set at time $t-1$, as in equation (9), we estimate the predictor coefficients of regression-based models and present the results in Table 3.

<Insert Table 3 about here>

In Panel A our results are similar to the findings of Hou et al. (2012), with future earnings being significantly positively correlated with total assets, dividend and lagged one-year earnings. We note that the average coefficient of lagged one-year earnings is highly significant, confirming the findings of Fama and French (2006) and Hou et al. (2012) that earnings are highly persistent and form the basis of most forecasts. Panel B shows the statistics of the HW model, which are in line with the results reported in Harris and Wang (2013).

Panel C presents the statistics of the KLSS model. These show that both changes in earnings analysts' forecasts and lagged one-year returns are positively and significantly correlated to the actual changes in report earnings. Interestingly, the adjusted-R^2 obtained from this regression is much lower than that of the HDZ model. However, this is not material because the KLSS model is based on changes in earnings as opposed to the level of earnings.

Finally, Panel D presents the statistics of the SO model. The explanatory power of the extended set of predictors (adjusted-R^2) is also lower than that of the HDZ model, but again it would be wrong to attach too much significance to the adjusted-R^2 because So (2013) is forecasting a different measure of earnings – an issue that we will have to deal with when we try to make comparisons between the competing models.

4. The forecasting models compared

4.1. Forecast bias of models

We next compute forecasts for each regression model by multiplying the estimated coefficients by the predictors at time t to generate earnings estimates for time $t+1$, as in equation (10). Earnings surprises are defined as actual reported earnings minus the forecasts, as in equation (11). We noted in Section 3 that our models aim to predict three different measures of earnings. The HDZ model predicts non-scaled earnings, while analysts and the SO and RW models predict EPS, and the KLSS model forecasts the future EPS/price ratio. To ensure the comparability of the models, we scale the different earnings surprises to produce a common measurement basis in the form of forward earnings yield (EPS/price).[11]

Table 4 presents the summary statistics of these earnings surprise proxies. Panel A compares the earnings surprises of alternative one-year-ahead earnings forecasts. The sizes and signs of the earnings surprises are indicative of any bias in the different earnings forecasts. As expected, by their construction, none of the regression-based models display statistically significant bias. In line with the literature, we observe a significant negative error in the forecasts made by analysts, indicating optimism. We should also remember that for comparability with our regression models, this statistic is for the earliest forecasts of analysts, which accounts for the relatively high degree of optimism (Ciciretti et al. 2009). Nevertheless, they are more accurate than forecasts of the RW, HDZ, HW and SO models based on the absolute earnings surprise (AES) statistics. Interestingly, there is no significant difference between the accuracy of the analysts' forecasts and the KLSS model.

<Insert Table 4 about here>

Nevertheless, like most previous studies, we have winsorized our data at the 1% level. Thus, we have effectively replaced 'unacceptable' actual reported earnings with a figure that

[11] The earnings surprises in the HDZ model are scaled by market capitalization, while those in the AF, RW, HW and SO models are scaled by adjusted stock prices to produce earnings surprises in the form of forward earnings yields. Since the KLSS model predicts changes in earnings per share deflated by price, there is no need to further scale earnings at this stage.

we regard as more acceptable for the fitting of the regression model. This puts analysts, who are allowed no such discretion, at a disadvantage in terms of forecast accuracy. We thus eliminate winsorized reported earnings data and run our subsequent analysis on the slightly reduced sample (Panel B). We find that in terms of accuracy as measured by the mean absolute deviation, analysts' forecasts are the most accurate.[12] This superiority is statistically significant (Panel C).[13] In addition, the variability in the absolute error of analysts' forecasts is the least. Among regression-based models, the KLSS model, which uses analysts' forecasts, past-year earnings and past-year stock return, outperforms the rest in terms of forecast accuracy.

In Panel D, we report the time-series average correlation of alternative proxies of earnings surprises. The correlations between all the models is positive. In general, the correlation coefficient reflects the degree of communality of the input dataset, with that of the KLSS model and analysts' forecasts being the highest. The KLSS model also has a high degree of correlation with the RW model. We will explore the implications of these observations in more depth in Section 5.2. The lower degree of correlation between analysts' forecasts with the regression models based purely on accounting values (HDZ, HW and SO) is a likely reflection of the fact that analysts also make extensive use of non-accounting information to form their predictions.

4.2. Earnings response coefficients and the preliminary result of the information theory

Our primary interest is the prediction of future price movements in response to errors in forecasts of earnings. Hence, we measure the earnings response coefficient (ERC), which reflects the stock market reaction for one unit of unexpected earnings by scaling all forecast

[12] The AESs in analysts' forecast reported are further disadvantaged by the 'optimism' of 0.0146 in Table 4. The triangle inequality implies that the absolute deviation is between the reported figure 0.0297 and 0.0151 (= 0.0297–0.0146) in the absence of such bias. In addition, the high value of the t-statistics suggests less variation in the absolute deviation than in the other measures.

[13] A negative entry in a column implies that the *column* variable has a lower absolute error, while a positive entry in a row implies that the *row* variable has a lower absolute error.

surprises by the standard deviation of surprises, using regression equation (12).[14] This structure emphasizes that ERCs vary across firms (j) and over time:

$$ACAR_{j,t+1} = a + ERC^{f} \times (e^{rep}_{j,t+1} - e^{f}_{j,t}) + u_{t+1} \qquad t = 0, T-1, \tag{12}$$

where $ACAR_{j,t+1}$ denotes annualized cumulative abnormal return which equals the sum of four quarterly earnings announcement abnormal returns (market adjusted, from day –1 to day 1).[15]

Table 5 presents the ERCs and their corresponding average t-statistics associated with each proxy of earnings surprises (Panel A), and a comparison between them (Panel B) obtained from the Fama–Macbeth two-stage regressions with Newey–West adjusted standard errors. The results in Panel A show that all the response coefficients are positive and statistically significant, with the ERCs of KLSS and analysts' forecasts being the highest. The adjusted-R^2 statistics are relatively low since we are predicting an 'earnings surprise' one year before it happens, and the relative uncertainty remaining in all the one-year-ahead forecasts is substantial.

<Insert Table 5 about here>

Panel B of Table 5 shows the differences between ERC estimates together with their corresponding t-statistics. We find that the ERCs of the KLSS model and analysts' forecasts are statistically indifferent yet they are statistically significantly higher than all the others. These are the best-performing models in terms of capturing the market expectation of future earnings. We also note in general the pivotal role of analysts' forecasts in forming market expectations, where it too dominates all the other models except for the KLSS model.

[14] Different models predict different measures of earnings, which prevents an easy comparison of ERCs. Hence, to enable a direct comparison, we follow Brown et al. (1987a) and Hou et al. (2012) by standardizing all proxies for earnings surprises each year to achieve unit variance. This has the added advantage in our case of making the resulting coefficients directly interpretable in terms of the theory developed in Section 2.

[15] We adopt the approach used by Hou et al. (2012) to estimate the unexpected returns (*UR*) in equations (4) and (5), which we refer to as $ACAR_{j,t+1}$ in equation (12).

We check the robustness of our findings by adding the contemporaneous values of control variables, which are used in the literature as the determinants of earnings response coefficients, into the return-earnings regression (Table 6). The control variable set includes the firms' beta (Collins and Kothari 1989, Zolotoy 2012), dividend pay-out ratio (Kallapur 1994, Mande 1994), leverage ratio, market-to-book ratio (MB) (Collins and Kothari 1989, Zolotoy 2012), negative earnings dummy and non-dividend paying dummy (Hayn 1995), earnings volatility (Dichev and Tang 2009), accrual quality (Perotti and Wagenhofer 2014) and the number of analysts' forecasts (Bartov et al. 1999).

We find that the changes in the numerical values of the ERCs of all the models are relatively minor. The ERC of the KLSS model remains the highest (from Panel A). Further tests, which are shown in Panel B of Table 6, indicate that it is statistically higher than those of the other forecasts, including analysts' forecasts. This confirms the robustness of earlier results. However, as we have shown in our theoretical framework, we still need to explore the relative contribution of modellers' forecasts in providing additional information about market expectations. We address this in the next section.

<Insert Table 6 about here>

4.3. Earnings surprise decomposition and the value of information

So far, we have investigated the properties of the forecasts in isolation. In general, we find that analysts' forecasts perform well in terms of forecast accuracy but present a systematic upward bias. Hence, to compare the relative merits of regression-based forecasts with those of analysts to ascertain whether predictable adjustments to analysts' forecasts exist, we decompose market earnings surprise into the two components outlined in Section 2 and perform the following regressions:

$$ACAR_{j,t+1} = \gamma_0 + \gamma_1 \times (e^{rep}_{j,t+1} - e^{a}_{j,t}) + \gamma_2 \times (e^{a}_{j,t} - e^{f}_{j,t}) + \mathrm{u}_{t+1} \quad t = 0, T-1, \tag{13}$$

where T is the number of years in our sample, $e^{rep}_{j,t+1} - e^{a}_{j,t}$ is the earnings surprise using analysts' one-year-ahead forecasts and $e^{a}_{j,t} - e^{f}_{j,t}$ is a measure of additional information contained in the regression model (i.e. the market's correction for analysts' forecast errors). Again, we report only aggregate time-series average values obtained from Fama–Macbeth two-stage regressions with Newey–West adjusted standard errors.

<Insert Table 7 about here>

Table 7 presents the time-series average of the earnings response coefficients associated with the earnings surprise arising from analysts' forecasts and adjustments for each of the models. We note that $\gamma_1 \geq \gamma_2$ is in line with our hypothesis in Section 2. We also note that the addition of the term $e^{rep}_{j,t+1} - e^{f}_{j,t}$ increases the explanatory power over that of analysts' forecasts alone. This suggests that all the 'mechanical' models including the RW contain additional information and provide corrections for the predictable component of the errors in the one-year-ahead analysts' forecasts (Hughes et al. 2008).

We rewrite equation (13) in the following form:

$$ACAR_{j,t+1} = \gamma_0 + (\gamma_1 - \gamma_2)(e^{rep}_{j,t+1} - e^{a}_{j,t}) + \gamma_2(e^{rep}_{j,t+1} - e^{f}_{j,t}) + \mathrm{u}_{t+1} \quad t = 0, T-1. \qquad (14)$$

As discussed in Section 2.2, in this form the significance of the difference $\gamma_1 - \gamma_2$ can be interpreted as a measure of the contribution of the analysts' forecast to the joint combination of forecasts, while the ratio $(\gamma_1 - \gamma_2) : \gamma_2$ is an indication of the relative weights to be attached to the two forecasts. First, we observe that the differences ($\gamma_1 - \gamma_2$) associated with the HDZ, SO, HW and RW models are all greater than 0.022 and that these differences are statistically significant. In contrast, that of the forecast produced by the KLSS model is much smaller (0.0025) and is not statistically significant, implying that analysts' forecasts carry no additional information to that already incorporated into the KLSS forecast. We also note that the relative weights as measured by the ratio $(\gamma_1 - \gamma_2) : \gamma_2$ to attach to the KLSS forecasts is of the order of 13 times that of the analysts' forecasts, whereas for all the other models the analysts' forecasts carry a weight greater than unity.[16] This confirms that the KLSS model is a better forecast of the market expectation of earnings expectations.

Finally, we conduct robustness checks by adding control variables in the return-earnings regression and present the time-series averages of its coefficients in Table 8. The relative performance of the models as measured by γ_1 and γ_2 remains largely unchanged, with the KLSS model being superior.

<Insert Table 8 about here>

5. Portfolio formation

5.1. Investment strategy

A principal reason for trying to get a better forecast of the market expectation of future earnings is to identify a strategy upon which we can construct a profitable investment portfolio.[17] In our case, we are dealing with one-year-ahead forecasts. If a subset of one-year-ahead forecasts proves to be overly pessimistic then we are likely to see an initial fall in value. Once actual earnings figures, which we assume will be higher than market expectations, are announced, stock prices are likely to respond positively. Hence, buying and holding stocks in this 'overly pessimistic' subset should prove a profitable investment strategy. On the other hand, for the subset of overoptimistic forecasts, overvaluation occurs – that is, buying and holding stocks in this subset leads to losses. The practical problem of these strategies is that the market expectation of future earnings is unobservable so the

[16] On checking, the correlation coefficients in Tables 7 and 8 and the observed values of the weights are found to be consistent with their theoretical relationships established in Appendix A (equation (A.7)).

[17] Our approach is distinct from that of So (2013) since we base our investment strategy on a proxy of expected market adjustment for analysts' forecast errors and mispredictions.

determination of whether particular forecasts are overly pessimistic is not immediately obvious.

The empirical results in Section 4 strongly suggest that earnings forecasts of the KLSS model are the best estimates of the market's expectation of one-year-ahead earnings while analysts' forecasts perform best in terms of forecast accuracy. We thus examine the difference between forecasts of the KLSS model ($e_{j,t}^{KLSS}$) and analysts ($e_{j,t}^{a}$). If $e_{j,t}^{KLSS} \ll e_{j,t}^{a}$, there is the possibility that the KLSS model or the market over-adjusts for those specific firms and produces an unduly pessimistic forecast, resulting in a series of positive earnings surprises. We therefore examine the excess abnormal returns (*EACAR*) associated with differences in model-based forecasts and analysts' earnings forecast differences $(e_{j,t}^{KLSS} - e_{j,t}^{a})$ (Table 9 and Figure 1). Here *EACAR* equals annualized cumulative abnormal returns (*ACAR*) minus the mean of annualized cumulative abnormal returns ($\overline{ACAR}$). We observe positive (negative) excess returns for the low (high) subset. This aligns with our prediction of a negative correlation between *EACAR* and the differences $(e_{j,t}^{KLSS} - e_{j,t}^{a})$. One possible explanation of this phenomenon of the difference in predictions between the one-year-ahead analysts' forecasts and the forecasts of the KLSS model is that the market takes more notice of the recent poor performance, which is captured in the KLSS model in preference to what is perceived to be new overly optimistic analysts' forecasts. Thus, at this early stage in the forecasting cycle, there is perhaps a degree of anchoring by the market (Tversky and Kahneman 1974, Shefrin 2000, Campbell and Sharpe 2009).

<Insert Table 9 about here>

In the next section, we simulate and evaluate the performance of an investment strategy of buying and holding stocks in the low (high) earnings forecast difference subsets.

5.2. Portfolio performance

We examine more formally an investment strategy that exploits the difference between the KLSS forecasts and analysts' forecasts. On 1st May each year, when all data and earnings forecasts are available for each year between 1985 and 2014, we first eliminate stocks in the top and bottom 1% of the earnings forecast differences to avoid picking up outliers. We then construct two buy-and-hold portfolios based on the differences in earnings forecasts of the KLSS model and those of analysts. One portfolio is formed from stocks in the lowest 10% (the low portfolio) and the other is formed from stocks in the highest 10% (the high portfolio).

We calculate buy-and-hold abnormal returns (BHAR) having adjusted for risks, as suggested by Fama (1998), using both the market model and the Fama–French five-factor model (Fama and French, 2015) with various holding periods ranging from a month to a year. Since we form an equally weighted portfolio, we use the equally weighted CRSP portfolio as a proxy for the market portfolio.

Panel A of Table 10 presents abnormal returns of the low portfolio where earnings forecasts of the KLSS model are significantly lower than those of analysts. We find that more than 70% of the time we get positive abnormal returns, and the mean of BHARs are also significantly positive regardless of the holding period. They are 3.8%, 8.3%, 11.7% and 13.8% for one-, three-, six- and twelve-month investments, respectively, if we use the market model (Panel A), ignoring transaction costs. If we use the Fama–French five-factor model, the corresponding results are 3.4%, 7.2%, 9.2% and 8.8%.[18] All are statistically significant.

<Insert Table 10 about here>

[18] When we apply the Fama–French (1993) three-factor model, the corresponding results are 3.7%, 7.9%, 10.9% and 12.36%.

As a further test of our model and theoretical explanation, we adopt the contra strategy of buying those shares where the forecasts of earnings produced by the KLSS are much higher than those of analysts – that is, $e_{j,t}^{KLSS} >> e_{j,t}^{a}$ (Panel B). As expected, here the earnings forecasts of the KLSS model are optimistic and the portfolio is subject to a series of negative earnings surprises. Losses occur in more than 93% of the years between 1985 and 2014.

If we examine the pattern of the excess returns, we notice that most of the gains in Table 10 Panel A are realized in a six-month holding period. These results are consistent with the explanation of a degree of anchoring suggested earlier in Section 5.1, whereby more weight is given to recent poor performance in the pessimistic set than to revised optimism by analysts. By the time the half-yearly results are announced, most of this early pessimism has been replaced by a more realistic view of the firm's performance. The reverse is true in the case of Table 10, Panel B, where poor performance is confirmed by the half-year stage and losses occur in the next six months of trading. This analysis confirms the greater ability of earnings forecast of the KLSS model to capture market sentiment in the form of expectations of one-year-ahead earnings, but equally does not deny the greater accuracy of analysts' forecasts.

6. Conclusion

In this paper, we develop a theoretical framework for comparing the information content of alternative earnings forecasts in terms of the correlation coefficients between errors in forecasts of market expectations. We establish conditions under which a forecast proxy dominates that of analysts and provides a sufficient estimate of the market expectation of future earnings. Based on the theory, we explore the information content of forecasts made by analysts and compare these with the RW model and four differently constructed regression-based approaches. We introduce a level playing field by removing any winsorized data, which is typically adjusted prior to regression analysis being carried out. We investigate the various forecasts in terms of forecast bias over actual reported earnings and the alignment to market expectation of future earnings.

We confirm that although analysts' forecasts are upwardly biased over actual reported earnings, in general within our subset of observations they are more accurate than other firm-characteristic regression-based models (the HDZ, HW and SO models). Meanwhile, the forecasts from the KLSS model, which in effect combines analysts' forecasts with lagged one-year-return adjustment, are less biased and have a higher earnings response coefficient. However, we also draw attention to the limitation of this and other studies. In our case, the need for a common dataset probably results in a concentration of larger companies with an established following of analysts. This does not deny the possibility that other regression models covering a greater number of smaller companies with a thinner analyst following could be found to outperform the one-year-ahead analysts' forecasts.

Finally, we conclude that within our dataset consisting of larger companies, the KLSS model still contains information additional to analysts' one-year-ahead forecasts. By concentrating on the most pessimistic subset, we can identify portfolios that outperform the market on a risk-adjusted basis. Further consideration of our results strongly suggests that the market initially forms its expectation based on recent performance, which forms part of the basis of the KLSS model. An application of this idea using US data from 1983 to 2015 suggests the existence of a profitable and exploitable investment strategy. We attribute this apparent violation of market efficiency to the market overweighting recent poor performance.

Our research draws attention to the importance of a theoretical framework that facilitates the understanding and interpretation of various empirical results. This analysis adds structure to the large literature on the superiority of various forecasting models. In our case, the statistical analysis of the underlying determinants of the earnings response coefficient provides the basis for comparing competing models used in forecasting earnings. Such models form the basis of equity valuation. It also provides a framework that helps understand how mechanistic forecasting models might outperform analysts' forecasts when analysts have access to sophisticated forecasting models, alongside superior private information. We argue that one answer lies in the market overweighting, or anchoring on, recent poor performance. We show how this observation can be exploited using a mechanistic selection technique that is not prone to such biases. This latter observation presents a rich vein for investment practitioners to explore and understand other similar contrarian strategies.

Appendix A: Decomposition of the maximum likelihood function of the market earnings surprises

As discussed in Section 2, we have the forecasting system following equation (2):

$$\begin{aligned} E^{rep} - E^{m} &= -\varepsilon_m \\ E^{rep} - E^{a} &= -\eta - \varepsilon_a \\ E^{rep} - E^{f} &= -\varepsilon_f \\ E^{a} - E^{f} &= \eta + \varepsilon_a - \varepsilon_f \end{aligned}$$.

Applying standard techniques for multivariate analysis from Press (1972 p. 69), we can deduce the theoretical values of the coefficients in equation (3) in terms of the variances and covariance of $\varepsilon_m, \varepsilon_a, \varepsilon_f$.

The maximum likelihood of the market earnings surprise given observed analysts' surprise and model-based surprise is as follows:

$$\mathrm{E}\{E^{rep} - E^{m} \mid E^{rep} - E^{a} = e^{rep} - \mathrm{e}^{a}, E^{a} - E^{f} = e^{a} - \mathrm{e}^{f}\} \sim N\left(\sum_{12}\sum_{22}^{-1}\begin{pmatrix} e^{rep} - \mathrm{e}^{a} + \eta \\ e^{a} - \mathrm{e}^{f} - \eta \end{pmatrix}, \sum_{11.2}\right)$$

(A.1)

where $\mathrm{E}\{.\}$ is the expectations operator, covariance matrices:

$$\Sigma_{12} = \left[\mathrm{cov}(\varepsilon_m, \varepsilon_a) \quad -\mathrm{cov}(\varepsilon_m, \varepsilon_a - \varepsilon_f)\right]$$

$$\Sigma_{22}^{-1} = \begin{bmatrix} \mathrm{var}(\varepsilon_a - \varepsilon_f) & \mathrm{cov}(\varepsilon_a, \varepsilon_a - \varepsilon_f) \\ \mathrm{cov}(\varepsilon_a, \varepsilon_a - \varepsilon_f) & \mathrm{var}(\varepsilon_a) \end{bmatrix} \det\left|\Sigma_{22}\right|^{-1}$$

and the variance term $\Sigma_{11.2} = \sigma^2(\varepsilon^m) - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. (A.2)

In practice, we standardize the observed values $e^{rep} - e^{a}$ and $e^{a} - e^{f}$ to have unit variance. If we do the same to the theoretical values, this simplification has the effect of simplifying the algebra. We have:

$$\Sigma_{12} = \left[\rho_{ma}\sigma_m \quad -\rho_{ma}\sigma_m + \rho_{mf}\sigma_m)\right] \text{ and } \Sigma_{22}^{-1} = \begin{bmatrix} 2(1-\rho_{af}) & 1-\rho_{af} \\ 1-\rho_{af} & 1 \end{bmatrix} \det\left|\Sigma_{22}\right|^{-1}, \quad \text{(A.3)}$$

where $\det\left|\Sigma_{22}\right|^{-1} = \dfrac{1}{1-\rho_{af}^2}$, ρ_{ma}, ρ_{mf} and ρ_{af} denote the correlation coefficient between ε_m and ε_a, ε_m and ε_f, ε_a and ε_f , respectively.

By multiplying the matrices out in equation (A.1), we get the coefficients in equation (3) in Section 2 as follows:

$$\lambda_1 = \frac{\rho_{ma}\sigma_m(1-\rho_{af}) + \rho_{mf}\sigma_m(1-\rho_{af})}{1-\rho_{af}^2} = \frac{(\rho_{ma}+\rho_{mf})}{1+\rho_{af}}\sigma_m > 0 \text{ if } \rho_{ma} > 0,\ \ \rho_{mf} > 0 \quad \text{(A.4)}$$

$$\lambda_2 = \frac{\sigma_m(\rho_{mf} - \rho_{ma}\rho_{af})}{1-\rho_{af}^2} \quad \text{(A.5)}$$

$$\lambda_1 - \lambda_2 = \frac{(\rho_{ma} - \rho_{mf}\rho_{af})\sigma_m}{1-\rho_{af}^2}. \quad \text{(A.6)}$$

This gives us a theoretical value for the relative weights in equation (3) as:

$$(\lambda_1 - \lambda_2) : \lambda_2 = (\rho_{ma} - \rho_{mf}\rho_{af}) : (\rho_{mf} - \rho_{ma}\rho_{af}). \quad \text{(A.7)}$$

We now consider the cases where the forecast produced by the regression model forecasts is a sufficient statistic for the analysts' forecast. We assume that $\varepsilon_a = \varepsilon_f + \varepsilon$, where ε is white noise $N(0,\sigma^2)$, and temporarily suspend the unit variance assumption ε_a and ε_f. Here analysts' forecasts are just a noisy version of the model's forecast.

Thus, $\mathrm{cov}(\varepsilon_m, \varepsilon) = \mathrm{cov}(\varepsilon_f, \varepsilon) = \mathrm{cov}(\varepsilon_m, \varepsilon_a - \varepsilon_f)\,\mathrm{cov}(\varepsilon_f, \varepsilon_a - \varepsilon_f) = 0$

$$\mathrm{cov}(\varepsilon_m, \varepsilon_a) = \rho_{ma}\sigma_m\sigma_a = \mathrm{cov}(\varepsilon_m, \varepsilon_f + \varepsilon) = \rho_{mf}\sigma_m\sigma_f$$

and $\mathrm{cov}(\varepsilon_a, \varepsilon_a - \varepsilon_f) = \mathrm{cov}(\varepsilon_a + \varepsilon_f + \varepsilon, \varepsilon) = \sigma^2$. (A.8)

Hence, $\Sigma_{12} = \left[\rho_{ma}\sigma_m\sigma_a \quad 0\right] = \left[\rho_{mf}\sigma_m\sigma_f \quad 0\right]$, $\Sigma_{22}^{-1} = \begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_a^2 \end{bmatrix} \det\left|\Sigma_{22}\right|^{-1}$

and $\lambda_1 = \lambda_2 = \dfrac{\rho_{mf}\sigma_m\sigma_f}{\sigma_a^2 - \sigma^2} = \dfrac{\rho_{mf}\sigma_m}{\sigma_f}$, (A.9)

which is just the regression coefficient of $e^{rep} - \mathrm{e}^m$ on $e^{rep} - \mathrm{e}^f$.

Appendix B: Variables' descriptions and information sets

Panel A: CRSP/COMPUSTAT items

Item	Description	Item	Description
AT	Total assets	TXP	Income taxes payable
DVC	Total dividends	IB	Income before extraordinary items
CSHO	Common shares outstanding	SPI	Special items
DP	Depreciation and amortization	CEQ	Total common/ordinary equity
ACT	Total current assets	CHE	Cash and short-term investment
LCT	Total current liabilities	OANCF	Operating activities net cash flow
DLC	Total debt in current liabilities	TCAP	Market capitalization

Panel B: Information set (IS_t) of regression-based models

HDZ model	Dependent variable: (IB) Net income before extraordinary items.
IB_t, AT_t, DVC_t	As above.
$DD1_t$	Indicator for dividend payer, which is one if dividends are positive and zero otherwise.
NEG_t	Negative earnings indicator, which is one if earnings are negative and zero otherwise.
ACR_t	Prior to 1988, accruals are the change in non-cash current assets less change in current liabilities, excluding the change in short-term debt and the change in taxes payable minus depreciation and amortization expenses. Since 1988, accruals equal the difference between earnings and cash flows from operations.
HW model	Dependent variable: (IBPS) net income before extraordinary items deflated by adjusted number of shares outstanding.
$IBPS_t$	Income before extraordinary items deflated by adjusted number of shares outstanding
B_t, B_{t-1}	Book value per share: book value deflated by adjusted number of shares outstanding.
$APRC_t$, $APRC_{t-1}$	Adjusted stock price three months after the fiscal year-end (end of March).
KLSS model	Dependent variable: (CIB) Actual change in earnings measured as changes in I/B/E/S EPS deflated by stock price.
RET_t	Lagged one-year compound returns from 1 April last year to 1 April this year.
CAF_t	Change in earnings forecasted by analysts, which is the analysts' median consensus forecast for the current period minus the previous period's earnings.
SO model	Dependent variable: (IBTS) Net income before extraordinary items and after subtracting special items and taxes deflated by adjusted number of shares outstanding.[19]
$IBTP_t$	Positive EPS, which equals EPS if they are positive and zero otherwise
$NEGE_t$	Negative earnings indicator, which is one if earnings are negative and zero otherwise.
ACP_t	Positive accruals equal accruals per share when accruals are positive and zero otherwise.[20]
ACN_t	Negative accruals equal accruals per share when accruals are negative and zero otherwise.
AG_t	Asset growth – percentage.
$DD2_t$	Indicator for non-dividend payer, one when firms do not pay dividend, and zero otherwise.
DIV_t	Dividends per share equal dividends for common equity divided by the number of shares.
BM_t	Book-to-market ratios equal common equity divided by market capitalization.
$APRC_t$	Adjusted stock price at fiscal year-end (end of December).

[19] This is to make the earnings figure from COMPUSTAT compatible with the I/B/E/S earnings per share (Bradshaw and Sloan 2002). The tax rate is assumed to be 35%.

[20] The accruals equal the change in current assets plus the change in debt in current liabilities minus the change in cash and short-term investments and minus the change in current liabilities.

Appendix B: Variables' descriptions and information sets (Cont.)

Other Variables	
Beta	Downloaded from WRDS/CRSP dated 11 July 2016.
DPR	Dividend pay-out ratios equal dividends divided by earnings before extraordinary items.
LEV	Leverage ratios which equal long-term debt divided by the market value of assets.
MB	Market-to-book ratios equal market capitalization divided by total common equity.
E_vol	Earnings volatility equals standard deviation of deflated earnings (total earnings/average total assets) of the last five years (minimum of three-year requirement).
DA	Discretionary accruals estimated the approach in Jones (1991).
NEAF	Number of analysts' forecasts from I/B/E/S.

References

Abarbanell, J. S. and Bushee, B. J. (1997). Fundamental analysis, future earnings, and stock prices. Journal of Accounting Research, 35 (1), 1–24.

Ashton, D. and Wang, P. G. (2013). Terminal valuations, growth rates and the implied cost of capital. Review of Accounting Studies, 18 (1), 261–290.

Ball, R. and P. Brown (1968). An empirical evaluation of accounting income numbers. Journal of Accounting Research, 159–178.

Bamber, L. S., Christensen, T. E. and Gaver, K. M. (2000). Do we really 'know' what we think we know? A case study of seminal research and its subsequent overgeneralization. Accounting, Organizations and Society, 25 (2), 103–129.

Bartov, E., Lynn, S. G. and Ronen, J. (1999). Returns-earnings regressions: an integrated approach. Available from: http://ssrn.com/abstract=163137 or http://dx.doi.org/10.2139/ssrn.163137.

Beaver, W. H. (1970). The time series behavior of earnings. Journal of Accounting Research, 8, 62–99.

Bosquet, K., de Goeij, P. and Smedts, K. (2015). Analysts' earnings forecasts: coexistence and dynamics of overconfidence and strategic incentives. Accounting and Business Research, 45(3), 307–322.

Botosan, C. A. and Plumlee, M. A. (2005). Assessing alternative proxies for the expected risk premium. The Accounting Review, 80 (1), 21–53.

Bradshaw, M. T. (2011). Analysts' forecast: what do we know after decades of research? Working Paper, Boston College. Available from: http://ssrn.com/abstract=1880339.

Bradshaw, M., Drake, M., Myers, J. and Myers, L. (2012). A re-examination of analysts' superiority over time-series forecasts of annual earnings. Review of Accounting Studies, 17 (4), 944–968. doi: 10.1007/s11142-012-9185-8.

Bradshaw, M. T., Richardson, S. A. and Sloan, R. G. (2001). Do analysts and auditors use information in accruals? Journal of Accounting Research, 39 (1), 45–74.

Bradshaw, M. T. and Sloan, R. G. (2002). GAAP versus the street: an empirical assessment of two alternative definitions of earnings. Journal of Accounting Research, 40 (1), 41–66.

Brown, L. D., Call, A. C., Clement, M. B. and Sharp, N. Y. (2015). Inside the 'black box' of sell-side financial analysts. Journal of Accounting Research, 53(1), 1–47.

Brown, L. D., Hagerman, R. L., Griffin, P. A. and Zmijewski, M. E. (1987a). An evaluation of alternative proxies for the market's assessment of unexpected earnings. Journal of Accounting and Economics, 9 (2), 159–193.

Brown, L. D., Hagerman, R. L., Griffin, P. A. and Zmijewski, M. E. (1987b). Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. Journal of Accounting and Economics, 9 (1), 61–87.

Brown, L. D. and Rozeff, M. S. (1978). The superiority of analyst forecasts as measures of expectations: evidence from earnings. The Journal of Finance, 33 (1), 1–16.

Campbell S. D. and Sharpe S. A. (2009). Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices. Journal of Financial and Quantitative Analysis, 44, 369–390.

Ciciretti, R., Dwyer, G. P., and Hasan, I. (2009). Investment analysts' forecasts of earnings. Federal Reserve Bank of St. Louis Review, 91 (5:2), 545–567.

Claus, J. and Thomas, J. (2001). Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. Journal of Finance, 56 (5), 1629–1666. doi: Doi 10.1111/0022-1082.00384.

Collins, D. W. and Kothari, S. (1989). An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. Journal of Accounting and Economics, 11 (2), 143–181.

Dichev, I. D. and Tang, V. W. (2009). Earnings volatility and earnings predictability. Journal of Accounting and Economics, 47 (1), 160–181.

Easton, P. D. and Monahan, S. J. (2005). An evaluation of accounting-based measures of expected returns. The Accounting Review, 80 (2), 501–538.

Easton, P. D. and Sommers, G. A. (2007). Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. Journal of Accounting Research, 45 (5), 983–1015.

Easton, P. D., Taylor, G., Shroff, P. and Sougiannis, T. (2002). Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. Journal of Accounting Research, 40 (3), 657–676. doi: Doi 10.1111/1475-679x.00066.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics, 49 (3), 283–306. doi: http://dx.doi.org/10.1016/S0304-405X(98)00026-9.

Fama, E. F. and K. R. French (1993). "Common risk factors in the returns on stocks and bonds." Journal of Financial Economics 33(1): 3-56.

Fama, E. F. and French, K. R. (2000). Forecasting profitability and earnings. The Journal of Business, 73 (2), 161–175.

Fama, E. F. and French, K. R. (2006). Profitability, investment and average returns. Journal of Financial Economics, 82 (3), 491–518.

Fama, E. F. and K. R. French (2015). "A five-factor asset pricing model." Journal of Financial Economics 116(1): 1-22.

Feldman, R., Lazer, R. and Livnat, J. (2003). Earnings guidance after regulation FD. Journal of Investing, Winter.

Foster, G. (1977). Quarterly accounting data: time-series properties and predictive-ability results. The Accounting Review, 52 (1), 1–21.

Frankel, R., Kothari, S. P. and Weber, J. (2006). Determinants of the informativeness of analyst research. Journal of Accounting and Economics 41 (1–2), 29–54.

Frankel, R. and Lee, C. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. Journal of Accounting and Economics, 25 (3), 283–319.

Fried, D. and Givoly, D. (1982). Financial analysts' forecasts of earnings: a better surrogate for market expectations. Journal of Accounting and Economics, 4 (2), 85–107.

Gebhardt, W. R., Lee, C. and Swaminathan, B. (2001). Toward an implied cost of capital. Journal of Accounting Research, 39 (1), 135–176.

Gerakos, J. and Gramacy, R. (2013). Regression-based earnings forecasts. Chicago Booth Research Paper No. 12–26. Available from: http://ssrn.com/abstract=2112137 or http://dx.doi.org/10.2139/ssrn.2112137.

Harris, R. D. F. and Wang, P. (2013). An improved earnings forecasting model. Available at: https://ssrn.com/abstract=2802019.

Hayn, C. (1995). The information content of losses. Journal of Accounting and Economics, 20 (2), 125–153.

Hou, K. W., van Dijk, M. A. and Zhang, Y. L. (2012). The implied cost of capital: a new approach. Journal of Accounting and Economics, 53 (3), 504–526. doi: DOI 10.1016/j.jacceco.2011.12.001.

Hughes, J., Liu, J. and Su, W. (2008). On the relation between predictable market returns and predictable analyst forecast errors. Review of Accounting Studies, 13 (2–3), 266–291.

Jones, J. (1991). Earnings Management During Import Relief Investigations. Journal of Accounting Research, 29, 193–228.

Kallapur, S. (1994). Dividend payout ratios as determinants of earnings response coefficients: a test of the free cash flow theory. Journal of Accounting and Economics, 17 (3), 359–375.

Konchitchki, Y., Lou, X., Sadka, G. and Sadka, R. (2013). Expected Earnings and the Post-Earnings-Announcement Drift. Available from: http://ssrn.com/abstract=1446929 or http://dx.doi.org/10.2139/ssrn.1446929.

Lee, W. Y., Jiang, C. X. and Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. Journal of Banking and Finance, 26, 2277–2299.

Li, K. and P. Mohanram (2014). Evaluating cross-sectional forecasting models for implied cost of capital. Review of Accounting Studies, 19(3), 1152–1185.

Mande, V. (1994). Earnings response coefficients and dividend policy parameters. Accounting and Business Research, 24 (94), 148–156.

O'Brien, P. C. (1988). Analysts' forecasts as earnings expectations. Journal of Accounting and Economics, 10 (1), 53–83.

Ohlson, A. (1991). The theory of value and earnings, and an introduction to the Ball–Brown analysis. Contemporary Accounting Research, 8 (1) 1–19.

Perotti, P. and Wagenhofer, A. (2014). Earnings quality measures and excess returns. Journal of Business Finance & Accounting, 41 (5–6), 545–571.

Press, S. J. (1972). Applied Multivariate Analysis. New York: Holt, Rinehart and Winston.

Rozeff, M. S. and Kinney, W. R. (1976). Capital market seasonality: the case of stock returns. Journal of Financial Economics, 3, 379–402.

Shefrin, H. (2000). Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing. Boston, MA: Harvard Business School Press.

Smith, D. and Pourciau, S. (1988). A comparison of the financial characteristics of December and non-December year-end companies. Journal of Accounting and Economics, 10, 335–344.

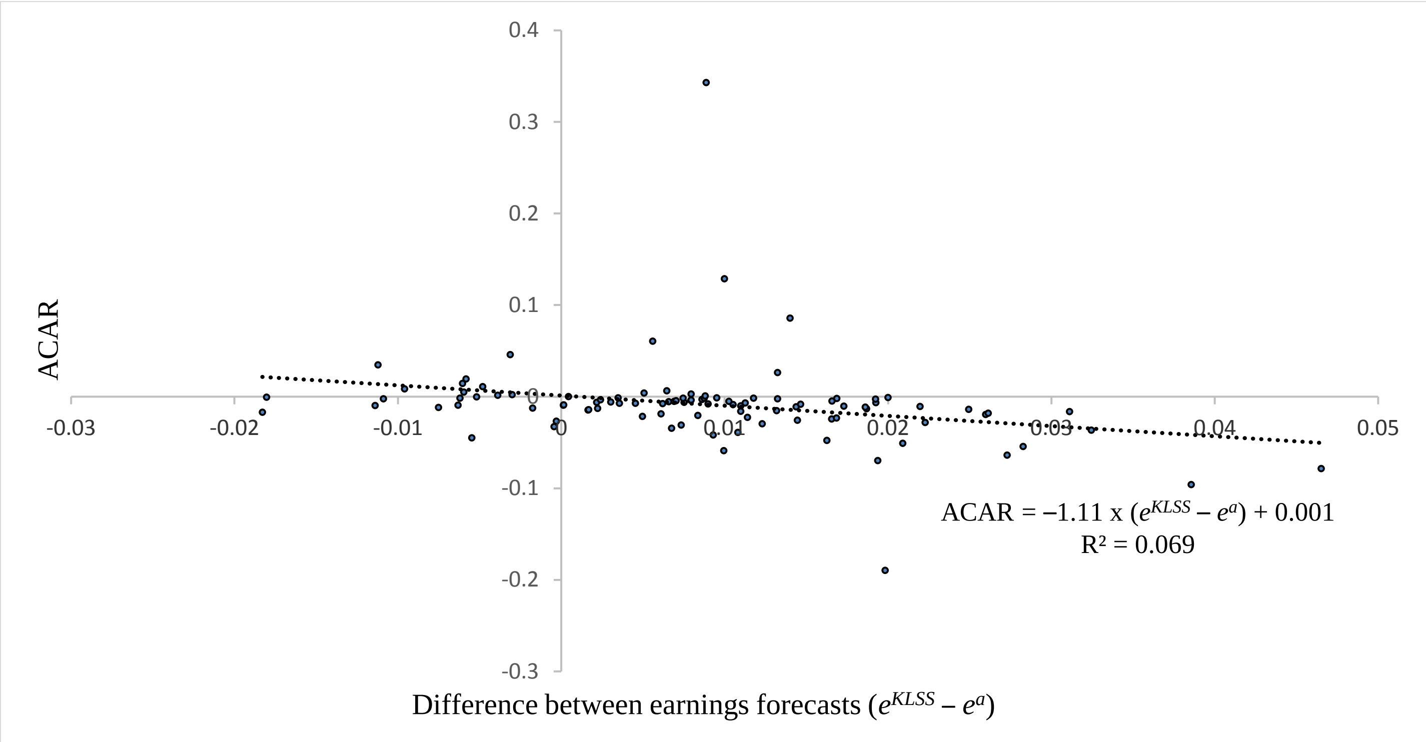
So, E. C. (2013). A new approach to predicting analyst forecast errors: do investors overweight analyst forecasts? Journal of Financial Economics, 108 (3), 615–640.

Tversky A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. Science, 185, 1124–1131.

Walther, B. R. (1997). Investor sophistication and market earnings expectations. Journal of Accounting Research, 35 (2), 157–179.

Zolotoy, L. (2012). Earnings surprise implicit in stock prices: which earnings forecasting models are investors using and what determines their choice? Journal of Business Finance & Accounting, 39 (9–10), 1161–1179. doi: DOI 10.1111/j.1468-5957.2012.02303.

Figure 1: The relationship between abnormal returns and differences in earnings forecasts



Annualized cumulative abnormal return (ACAR) plotted against the difference in forecast future earnings scaled by stock price between those made by the KLSS model (e^{KLSS}) and analysts (e^{a}). The data consists of 26,500 individual observations across all years ranked on the differences, with each data point representing the average values of portfolios of size 265.

Table 1: Data and number of observations

This presents the number of observations at each stage of our forecasting and evaluation exercise.

Data scope	December fiscal year-end firms with sharecode 10/11[a]					
Time	1983–2015 for accounting data and 1983–2015 for stock data file					
Model	Total[b]	(1)[c]	(2)[d]	(3)[e]	(4)[f]	(5)[g]
AF	59442	59442	29469	27906	27903	26506
RW	63651	63651				
HDZ	97283	84271				
HW	81189	79270				
KLSS	52981	47907				
SO	73825	61679				
ACAR						

[a] Firms in NYSE, NASDAQ and AMEX stock exchanges.

[b] To ensure the matching window between returns and earnings surprises and align firms' business nature, our study uses only firms with December fiscal year-end and sharecode 10/11 (excluding ADRs, close-end funds and REITS), which account for about 55% of the whole population.

[c] Earnings forecasts: to generate earnings forecasts, we need information available for at least two consecutive years (for the case of the HDZ, SO and KLSS models) or three consecutive years (for the case of the HW model). This reduces the number of firm-year observations of earnings forecasts.

[d] Number of earnings forecast observations available for all models.

[e] Scaling earnings surprises: earnings surprises in year $t+1$ are scaled by adjusted share prices in March of year *t* for the case of the AF, RW, HW and SO models and by market capitalization in March of year t for the case of the HDZ model. The absence of price information causes a slight fall in the number of firm-year observations.

[f] Matching earnings surprise return: returns (*ACAR*) are the annualized cumulative abnormal returns of four quarterly earnings announcements that are market adjusted from day –1 to day +1. We require firms to have non-missing price information for those days. In addition, there are cases where the earnings forecast is available while the actual reported earnings are missing. The absence of price information or of the actual reported earnings for some firms in some years causes a fall in the number of observations when we match earnings surprises with the returns.

[g] Excluding winsorized earnings for a fair forecast accuracy comparison.

Table 2: Summary statistics of the variables in the cross-sectional earnings models (1983–2014)

This provides summary time-series average statistics (mean, median, standard deviation and selected percentiles) of the variables used in the regression-based models.

	Description	Mean	P1	Median	P99	SD
Panel A: Variables in HDZ model (Hou et al. 2012)						
IB	Income before extraordinary items ($m)	86.1	–400.7	5.2	2232.6	335.7
AT	Total assets ($m)	3077.2	3.9	348.7	62830.7	9176.6
DVC	Total dividends ($m)	32.55	0.00	0.03	768.14	114.28
DD1	Indicator for dividend payer	0.38	0.00	0.09	1.00	0.48
NEG	Negative earnings indicator	0.34	0.00	0.03	1.00	0.47
ACR	Accruals ($m)	–114.6	–2318.7	–10.3	207.3	354.2
Number of observations:		97283				
Panel B: Variables in HW model (Harris and Wang 2013)						
IBPS	EPS ($)	0.30	–17.81	0.48	7.64	3.08
B	Book value per share ($)	10.36	–3.97	6.55	89.99	14.07
APRC	Adjusted stock price (end March) ($)	19.45	0.73	12.46	169.82	25.68
Number of observations:		81189				
Panel C: Variables in KLSS model (Konchitchki et al. 2013)						
CIB	Actual change in EPS deflated by stock prices	0.004	–0.380	0.006	0.455	0.094
RET	Lagged one-year returns	0.197	–0.636	0.106	2.291	0.614
CAF	Change in EPS forecasted by analysts	0.024	–0.120	0.009	0.487	0.074
Number of observations:		52981				
Panel D: Variables in SO model (So 2013)						
IBTS	EPS adjusted for special items ($)	0.22	–8.51	0.25	4.29	1.54
IBTP	Positive EPS ($)	0.56	0.00	0.25	4.29	0.80
NEGE	Negative earnings indicator ($)	0.33	0.00	0.00	1.00	0.47
ACP	Positive accruals per share ($)	0.42	0.00	0.03	6.55	0.98
ACN	Negative accruals per share ($)	0.37	0.00	0.01	7.04	1.01
AG	Asset growth – percentage	7.57	–67.38	3.84	178.27	34.72
DD2	Indicator for non-dividend payer	0.66	0.00	0.91	1.00	0.46
DIV	Dividends per share ($)	0.21	0.00	0.00	2.41	0.45
BM	Book-to-market ratio	0.60	–1.90	0.51	3.37	0.66
APRC	Adjusted stock price (end December) ($)	17.94	0.37	11.19	152.64	22.52
Number of observations:		73825				

Table 3: Earnings regression coefficients (1984–2014)

This presents the time-series average coefficients and their t-statistics from cross-sectional regression of earnings (e_t^{rep}) on the lagged-one-year of information set ($\mathbf{IS}_{j,t-1}$): $e_t^{rep} = \alpha + \boldsymbol{\beta} \times \mathbf{IS}_{t-1} + \varepsilon_t$. Detailed descriptions of dependent and independent variables are shown in Appendix B.

Variables	Description	Average coefficient	T-statistic
Panel A: Earnings regressions using HDZ model			
IB	Income before extraordinary items	0.721	*(26.22)*
AT	Total assets	0.003	*(2.50)*
DVC	Total dividends	0.363	*(7.43)*
DD1	Indicator for dividend payer	–0.542	*(–0.24)*
NEG	Negative earnings indicator	1.923	*(0.87)*
ACR	Accruals	–0.079	*(–4.40)*
Intercept		–3.618	*(–2.99)*
Adj. R^2		0.785	
Panel B: Earnings regressions using HW model			
IBPS	EPS before extraordinary items	0.610	*(22.31)*
B	Book value per share	–0.003	*(–0.33)*
LB	Lagged one-year book value	0.008	*(1.34)*
APRC	Adjusted stock price (end March)	0.010	*(2.71)*
LAPRC	Lagged one-year adjusted price (end March)	–0.018	*(–5.79)*
Intercept		0.209	*(4.67)*
Adj. R^2		0.4235	
Panel C: Earnings regressions using KLSS model			
RET	Lagged one-year returns	0.0304	*(4.17)*
CAF	Change in EPS forecasted by analysts	0.7495	*(21.28)*
Intercept		–0.0154	*(–4.06)*
Adj. R^2		0.377	
Panel D: Earnings regressions using SO model			
IBTP	Positive EPS	1.159	*(21.05)*
NEGE	Negative earnings indicator	–0.541	*(–10.66)*
ACP	Positive accruals per share	–0.148	*(–7.08)*
ACN	Negative accruals per share	–0.177	*(–8.41)*
AG	Asset growth – percentage	0.001	*(2.31)*
DD2	Indicator for non-dividend payer	–0.097	*(–7.11)*
DIV	Dividends per share	0.329	*(5.65)*
BM	Book-to-market ratio	–0.120	*(–3.79)*
APRC	Adjusted stock price (end December)	–0.024	*(–5.90)*
Intercept		0.364	*(9.40)*
Adj. R^2		0.497	

Table 4: Properties of earnings surprises and forecast errors (1985–2014)

This presents summary statistics of earnings surprise that equal actual earnings minus earnings forecasts (scaled by the market capitalization of the previous period for the HDZ model and by the stock price of the previous period for the AF, RW, HW and SO models). The results are based on the sample of firm-year observations for which forecasts are available for all models. Panel A (B) shows the time-series averages of the means of the earnings surprises and of the AESs and their t-statistics when we include (exclude) winsorized earnings in the test. Panel C shows the statistical significance of the differences in absolute errors of alternative earnings forecasts presented in Panel B. Panel D shows the time-series averages of correlations of the alternative proxies of earnings surprises.

Panel A: Including winsorized earnings

		AF	RW	HDZ	HW	KLSS	SO
Mean ES		–0.0215	0.0090	0.0116	–0.0056	–0.0036	0.0031
	t-stat	*(–4.26)*	*(3.00)*	*(0.77)*	*(–0.63)*	*(–0.98)*	*(1.08)*
Mean AES		0.0391	0.0527	0.0933	0.0835	0.0405	0.0493
	SD	0.0046	0.0062	0.0194	0.0132	0.0055	0.0058
No.		27903	27903	27903	27903	27903	27903

Panel B: Excluding winsorized earnings

		AF	RW	HDZ	HW	KLSS	SO
Mean ES		–0.0146	0.0040	0.0110	–0.0030	–0.0015	0.0042
	t-stat	*(–5.14)*	*(2.17)*	*(0.94)*	*(–0.44)*	*(–0.41)*	*(1.54)*
Mean AES		0.0297	0.0376	0.0824	0.072	0.0367	0.045
	SD	0.0026	0.0038	0.0160	0.0117	0.0052	0.0053
No.		26506	26506	26506	26506	26506	26506

Panel C: Differences between AESs observed in Panel B

		AF	RW	HDZ	HW	KLSS	SO
RW		–0.0079					
	t-stat	*(–4.08)*					
HDZ		–0.0528	–0.0448				
	t-stat	*(–3.61)*	*(–3.49)*				
HW		–0.0424	–0.0344	0.0104			
	t-stat	*(–4.16)*	*(–4.11)*	*(2.16)*			
KLSS		–0.0071	0.0009	0.0457	0.0353		
	t-stat	*(–2.03)*	*(0.37)*	*(3.60)*	*(4.25)*		
SO		–0.0153	–0.0074	0.0374	0.0270	–0.0083	
	t-stat	*(–3.81)*	*(–3.05)*	*(3.40)*	*(3.97)*	*(–2.34)*	

Panel D: Correlations of earnings surprises across models

	AF	RW	HDZ	HW	KLSS	SO
RW	0.612					
HDZ	0.439	0.424				
HW	0.426	0.425	0.645			
KLSS	0.811	0.709	0.419	0.416		
SO	0.354	0.427	0.402	0.497	0.364	

Table 5: Earnings response coefficients (1985–2014)

This presents ERCs associated with all proxies of earnings surprises and an analysis of the differences between them. These results are based on the 26,506 firm-year observations that are available for all models. Panel A illustrates the time-series averages of ERCs and their t-statistics obtained from the cross-sectional regression of annualized cumulative abnormal earnings announcement returns $(ACAR_{j,t+1})$ on earnings surprise estimates $(e^{rep}_{j,t+1} - e^{f}_{j,t})$:

$$ACAR_{j,t+1} = a + ERC^{f}_{t+1} \times (e^{rep}_{j,t+1} - e^{f}_{j,t}) + u_{t+1} \qquad t = 0, T-1 .$$

Panel B shows the time-series averages of the differences between ERC estimates together with the corresponding t-statistics in parentheses.

Panel A: ERCs						
Variable	AF	RW	HDZ	HW	KLSS	SO
ERCs	0.0317	0.0285	0.02	0.018	0.0339	0.0142
t-stat	*(10.67)*	*(13.50)*	*(10.15)*	*(7.06)*	*(12.43)*	*(10.59)*
Intercepts	0.0145	0.0053	0.0084	0.0092	0.0079	0.0072
t-stat	*(4.66)*	*(1.70)*	*(3.77)*	*(3.51)*	*(3.20)*	*(2.74)*
Adj. R^2	0.0428	0.0351	0.0189	0.015	0.0493	0.0108
N	26506	26506	26506	26506	26506	26506

Panel B: Differences between ERC estimates						
	AF	RW	HDZ	HW	KLSS	SO
RW	0.0032					
t-stat	*(2.15)*					
HDZ	0.0117	0.0084				
t-stat	*(6.66)*	*(5.27)*				
HW	0.0136	0.0104	0.002			
t-stat	*(8.98)*	*(5.47)*	*(1.20)*			
KLSS	–0.0022	–0.0055	–0.0139	–0.0159		
t-stat	*(–1.57)*	*(–3.61)*	*(–7.76)*	*(–7.83)*		
SO	0.0175	0.0142	0.0058	0.0038	0.0169	
t-stat	*(8.70)*	*(9.21)*	*(3.27)*	*(1.86)*	*(7.19)*	

Table 6: Earnings response coefficients with control variables (1985–2014)

This presents the time-series averages of ERCs and their t-statistics obtained from the cross-sectional regression of annualized cumulative abnormal earnings announcement returns ($ACAR_{j,t+1}$) on earnings surprise estimates ($e^{rep}_{j,t+1} - e^{f}_{j,t}$) and control variables including earnings volatility (E_vol), discretionary accruals (DA), leverage (Lev), market-to-book ratio (MB), beta, number of analysts' forecasts (NEAF), negative earnings dummy (NEG), dividend paying dummy (DD1) and dividend pay-out ratios (DPR):

$$ACAR_{j,t+1} = a + ERC^{f}_{t+1} \times (e^{rep}_{j,t+1} - e^{f}_{j,t}) + \sum_{i=1}^{j} \beta_{i,t+1} control_{i,t+1} + u_{t+1}, \qquad t = 0, T-1.$$

Panel A: ERCs with control variables						
	AF	RW	HDZ	HW	KLSS	SO
ERC	0.0285	0.0253	0.0148	0.013	0.0306	0.0139
t-stat	*(13.08)*	*(12.66)*	*(9.64)*	*(8.05)*	*(13.72)*	*(8.76)*
E_vol	0.007	–0.0192	–0.0012	0.0005	0.0049	0.0016
t-stat	*(0.22)*	*(–0.56)*	*(–0.04)*	*(0.02)*	*(0.16)*	*(0.05)*
DA	–0.0617	–0.0589	–0.0695	–0.0651	–0.0594	–0.0591
t-stat	*(–3.00)*	*(–2.80)*	*(–3.17)*	*(–2.92)*	*(–3.09)*	*(–2.77)*
Lev	0.0164	0.0107	0.0135	0.0109	0.0145	0.0132
t-stat	*(3.26)*	*(1.96)*	*(2.57)*	*(2.15)*	*(2.89)*	*(2.48)*
MB	–0.0004	–0.0004	–0.0005	–0.0004	–0.0003	–0.0004
t-stat	*(–2.37)*	*(–1.94)*	*(–2.39)*	*(–2.34)*	*(–1.88)*	*(–2.21)*
Beta	–0.0013	0.0009	0.0002	–0.0005	0.0007	0.0002
t-stat	*(–0.41)*	*(0.25)*	*(0.08)*	*(–0.14)*	*(0.22)*	*(0.07)*
NEAF	–0.0004	–0.0001	–0.0001	–0.0002	–0.0004	–0.0002
t-stat	*(–2.41)*	*(–0.59)*	*(–0.85)*	*(–1.44)*	*(–2.64)*	*(–1.07)*
NEG	–0.0383	–0.0485	–0.0517	–0.0536	–0.0396	–0.0628
t-stat	*(–5.67)*	*(–6.98)*	*(–8.20)*	*(–9.88)*	*(–5.72)*	*(–10.37)*
DD1	–0.0188	–0.0149	–0.0155	–0.0194	–0.0181	–0.0156
t-stat	*(–5.41)*	*(–4.62)*	*(–4.92)*	*(–5.64)*	*(–5.36)*	*(–4.47)*
DPR	–0.002	–0.0022	–0.0024	–0.0027	–0.002	–0.0028
t-stat	*(–2.27)*	*(–2.58)*	*(–2.41)*	*(–2.45)*	*(–2.30)*	*(–2.68)*
Intercept	0.0347	0.0235	0.0269	0.0314	0.0268	0.0282
t-stat	*(6.51)*	*(4.88)*	*(5.97)*	*(6.20)*	*(5.79)*	*(5.57)*
Adj. R^2	0.0731	0.0704	0.0529	0.0509	0.0787	0.0539
N	25911	25911	25911	25911	25911	25911

Panel B: Differences between ERC estimates						
	AF	RW	HDZ	HW	KLSS	SO
RW	0.0032					
t-stat	*(2.13)*					
HDZ	0.0137	0.0105				
t-stat	*(7.43)*	*(5.82)*				
HW	0.0155	0.0124	0.0018			
t-stat	*(11.39)*	*(6.48)*	*(1.11)*			
KLSS	–0.0021	–0.0052	–0.0158	–0.0176		
t-stat	*(–1.72)*	*(–3.91)*	*(–8.81)*	*(–10.05)*		
SO	0.0146	0.0114	0.0009	–0.001	0.0166	
t-stat	*(9.08)*	*(7.82)*	*(0.49)*	*(–0.52)*	*(8.82)*	

Table 7: Earnings surprise decomposition and earnings response coefficients (1985–2014)

This presents the time-series averages of ERCs associated with the earnings surprise arising from analysts' forecasts ($e^{rep}_{j,t+1} - e^{a}_{j,t}$) and the model-based market's expected correction for analysts' errors, which equals the standardized earnings surprises of the model-based forecasts minus the standardized earnings surprises of analysts' forecasts ($e^{a}_{j,t} - e^{f}_{j,t}$) and the corresponding t-statistics from the regressions:

$$ACAR_{j,t+1} = \gamma_{0,t+1} + \gamma_1(e^{rep}_{j,t+1} - e^{a}_{j,t}) + \gamma_2(e^{a}_{j,t} - e^{f}_{j,t}) + \mathrm{u}_{t+1} \qquad t = 0, T-1.$$

		AF	RW	HDZ	HW	KLSS	SO
γ_1		0.0317	0.0386	0.0392	0.0367	0.0353	0.0385
	t-stat	*(10.67)*	*(12.40)*	*(9.62)*	*(8.70)*	*(12.49)*	*(11.57)*
γ_2			0.0158	0.011	0.0075	0.0328	0.0084
	t-stat		*(8.73)*	*(5.86)*	*(3.40)*	*(8.69)*	*(8.40)*
Intercepts		0.0145	0.0110	0.0135	0.0144	0.0076	0.0137
	t-stat	*(4.66)*	*(3.63)*	*(4.91)*	*(4.78)*	*(3.56)*	*(4.48)*
Adj. R^2		0.0428	0.0525	0.0485	0.0467	0.0535	0.0474
N		26506	26506	26506	26506	26506	26506
$\gamma_1 - \gamma_2$			0.0228	0.0282	0.0292	0.0025	0.0301
	t-stat		*(9.52)*	*(12.81)*	*(15.38)*	*(0.71)*	*(14.33)*
$(\gamma_1 - \gamma_2) : \gamma_2$			1.44:1	2.56:1	3.89:1	0.08:1	3.58:1

Table 8: Earnings surprise decomposition and earnings response coefficients with control variables (1985–2014)

This presents the time-series averages of ERCs associated with the earnings surprise arising from analysts' forecasts ($e_{t+1}^{rep} - e_t^a$), the expected market's correction for analysts' errors ($e_t^a - e_t^f$) and control variables including earnings volatility (E_vol), discretionary accruals (DA), leverage (Lev), market-to-book ratio (MB), beta, number of analysts' forecasts (NEAF), negative earnings dummy (NEG), dividend paying dummy (DD1) and dividend pay-out ratios (DPR):

$$ACAR_{j,t+1} = \gamma_0 + \gamma_1(e_{j,t+1}^{rep} - e_{j,t}^a) + \gamma_2(e_{j,t}^a - e_{j,t}^f) + \sum_{i=1}^{j} \beta_i control_{i,t+1} + \mathrm{u}_{t+1} \quad t = 0, T-1.$$

		AF	RW	HDZ	HW	KLSS	SO
γ_1		0.0285	0.0352	0.0352	0.0328	0.0318	0.0357
	t-stat	*(13.08)*	*(13.53)*	*(11.60)*	*(11.15)*	*(14.79)*	*(12.46)*
γ_2			0.0153	0.0084	0.0055	0.0315	0.009
	t-stat		*8.3534*	*4.5519*	*4.0411*	*7.9535*	*6.692*
E_vol		0.007	–0.0114	–0.0012	0.0025	0.0027	0.0015
	t-stat	*(0.22)*	*(–0.32)*	*(–0.04)*	*(0.08)*	*(0.08)*	*(0.05)*
DA		–0.0617	–0.0625	–0.0718	–0.0673	–0.0589	–0.0667
	t-stat	*(–3.00)*	*(–3.13)*	*(–3.64)*	*(–3.29)*	*(–3.10)*	*(–3.39)*
Lev		0.0164	0.0137	0.0166	0.0157	0.0148	0.0165
	t-stat	*(3.26)*	*(2.62)*	*(3.10)*	*(3.12)*	*(2.91)*	*(3.15)*
MB		–0.0004	–0.0004	–0.0004	–0.0004	–0.0003	–0.0004
	t-stat	*(–2.37)*	*(–2.23)*	*(–2.38)*	*(–2.35)*	*(–1.79)*	*(–2.29)*
Beta		–0.0013	–0.0003	–0.0009	–0.0015	0.0007	–0.0011
	t-stat	*(–0.41)*	*(–0.10)*	*(–0.28)*	*(–0.48)*	*(0.21)*	*(–0.32)*
NEAF		–0.0004	–0.0003	–0.0003	–0.0004	–0.0004	–0.0004
	t-stat	*(–2.41)*	*(–1.76)*	*(–2.17)*	*(–2.44)*	*(–2.58)*	*(–2.41)*
NEG		–0.0383	–0.0361	–0.0336	–0.035	–0.0374	–0.0387
	t-stat	*(–5.67)*	*(–4.96)*	*(–4.89)*	*(–5.48)*	*(–5.32)*	*(–5.60)*
DD1		–0.0188	–0.0162	–0.0169	–0.0189	–0.0179	–0.0169
	t-stat	*(–5.41)*	*(–4.85)*	*(–5.10)*	*(–5.46)*	*(–5.27)*	*(–4.74)*
DPR		–0.002	–0.0018	–0.0018	–0.002	–0.002	–0.0019
	t-stat	*(–2.27)*	*(–2.33)*	*(–2.12)*	*(–2.19)*	*(–2.28)*	*(–2.28)*
Intercept		0.0347	0.0289	0.0316	0.0345	0.0262	0.0328
	t-stat	*(6.51)*	*(5.86)*	*(6.36)*	*(6.39)*	*(5.93)*	*(6.08)*
Adj. R^2		0.0731	0.0819	0.0768	0.0751	0.0822	0.0779
N		25911	25911	25911	25911	25911	25911
$\gamma_1 - \gamma_2$			0.0199	0.0268	0.0272	0.0003	0.0267
	t-stat		*(10.21)*	*(14.16)*	*(16.11)*	*(0.07)*	*(15.46)*
$(\gamma_1 - \gamma_2):\gamma_2$			1.30:1	3.19:1	4.96:1	0.01:1	2.97:1

Table 9: Abnormal returns and earnings forecast differences

This shows the average abnormal returns for each decile of the differences between forecasts of the KLSS model and analysts as measured by the forward earnings yield forecasts. $ACAR$ denotes annualized cumulative abnormal earnings announcement returns, which equal the sum of four quarterly earnings announcement abnormal returns. $EACAR$ denotes excess abnormal returns, which equal annualized cumulative abnormal returns minus the mean of annualized cumulative abnormal returns $(EACAR = ACAR - \overline{ACAR})$.

Decile portfolio		No.	$e^{KLSS} - e^{a}$	ACAR	EACAR
D1		2650	–0.0768	0.0219	0.0133
	t-stat		*(–61.18)*	*(5.92)*	*(3.59)*
D2		2651	–0.0325	0.0114	0.0028
	t-stat		*(–330)*	*(3.32)*	*(0.82)*
D3		2650	–0.0202	0.0131	0.0045
	t-stat		*(–400)*	*(3.91)*	*(1.33)*
D4		2651	–0.0135	0.0082	–0.0004
	t-stat		*(–500.00)*	*(2.33)*	*(–0.11)*
D5		2651	–0.0096	0.0073	–0.0013
	t-stat		*(–510.00)*	*(2.38)*	*(–0.43)*
D6		2650	–0.0066	0.0073	–0.0013
	t-stat		*(–440)*	*(2.43)*	*(–0.42)*
D7		2651	–0.0038	0.0101	0.0015
	t-stat		*(–220)*	*(3.33)*	*(0.50)*
D8		2650	–0.0014	0.0042	–0.0044
	t-stat		*(–170)*	*(1.41)*	*(–1.48)*
D9		2651	0.0032	0.0003	–0.0083
	t-stat		*(59.45)*	*(0.08)*	*(–2.43)*
D10		2651	0.0800	0.0022	–0.0064
	t-stat		*(28.06)*	*(0.61)*	*(–1.74)*
Mean			–0.0081	0.0086	0
	t-stat		*(–21.12)*	*(8.16)*	

Table 10: Buy-and-hold abnormal returns

This presents abnormal returns obtained from buying and holding stocks in the bottom (low) and the top (high) ten percentiles of the differences between earnings forecasts of the KLSS model and analysts each year from 1985 to 2014. Stocks are picked at the end of April each year based on the differences between forecasts of one-year-ahead earnings of the KLSS model (e^{KLSS}) and analysts (e^{a}) and then buy and hold for different investment periods.

Panel A: Portfolios for which KLSS forecasts are much more pessimistic $(e^{KLSS} \ll e^{a})$

		Market model				Fama–French five-factor model			
Year	Size	(1)[a]	(2)	(3)	(4)	(1)	(2)	(3)	(4)
1985	45	0.038	0.122	0.217	0.098	0.028	0.095	0.155	0.088
1986	46	0.008	0.025	0.167	0.263	0.012	0.040	0.169	0.303
1987	42	0.053	0.094	0.107	0.086	0.050	0.095	0.116	0.035
1988	42	0.003	0.020	0.016	0.034	0.005	0.023	0.004	–0.007
1989	45	0.036	0.079	0.078	0.143	0.032	0.074	0.067	0.100
1990	48	0.040	0.121	0.112	0.091	0.032	0.126	0.120	0.163
1991	49	0.034	0.126	0.056	–0.100	0.028	0.123	0.036	–0.023
1992	54	0.034	0.017	0.000	–0.298	0.025	0.014	0.003	–0.213
1993	59	0.061	0.109	0.154	0.148	0.033	0.053	0.037	–0.295
1994	65	0.022	0.016	0.048	–0.046	0.021	0.004	–0.031	–0.298
1995	69	0.036	–0.045	–0.114	–0.242	0.034	–0.036	–0.110	–0.242
1996	83	0.005	0.052	0.057	0.140	0.005	0.044	0.035	0.102
1997	88	0.027	–0.003	–0.013	–0.202	0.014	–0.005	–0.016	–0.215
1998	94	0.028	0.043	0.094	–0.008	0.034	0.006	0.008	–0.216
1999	95	0.042	0.159	0.286	0.959	0.041	0.149	0.257	0.883
2000	93	0.049	0.059	0.149	0.245	0.057	0.053	0.120	0.005
2001	89	0.007	0.050	–0.001	–0.191	0.010	0.052	0.032	–0.153
2002	95	0.018	0.030	0.051	–0.054	0.027	0.021	0.020	–0.127
2003	105	0.102	0.259	0.488	0.542	0.100	0.256	0.516	0.619
2004	109	0.049	0.159	0.214	0.398	0.022	0.087	0.121	0.279
2005	108	–0.001	0.051	0.084	0.064	–0.003	0.058	0.071	0.096
2006	109	0.028	0.060	0.130	0.245	0.036	0.090	0.157	0.290
2007	111	0.026	0.055	0.043	0.042	0.032	0.037	–0.002	0.011
2008	113	0.032	0.124	0.258	0.233	0.020	0.079	0.219	0.199
2009	120	0.188	0.320	0.308	0.434	0.193	0.319	0.322	0.378
2010	127	0.045	0.124	0.194	0.463	0.016	0.059	0.078	0.247
2011	122	0.017	0.013	–0.025	–0.028	0.020	0.023	–0.026	–0.003
2012	121	0.041	0.056	0.113	0.235	0.032	0.052	0.108	0.244
2013	124	0.074	0.151	0.161	0.338	0.058	0.119	0.109	0.306
2014	128	0.004	0.039	0.075	0.110	0.013	0.043	0.064	0.095
Average		0.038	0.083	0.117	0.138	0.034	0.072	0.092	0.088
t-stat		*(5.83)*	*(5.95)*	*(5.38)*	*(2.91)*	*(5.24)*	*(5.40)*	*(4.15)*	*(1.82)*
Positive		97%	93%	83%	70%	97%	93%	83%	63%
Negative		3%	7%	17%	30%	3%	7%	17%	37%

[a] (1) one-month holding. (2) three-month holding. (3) six-month holding. (4) one-year holding.

Table 10: Buy-and-hold abnormal returns (Cont.)

Panel B: Portfolios for which KLSS forecasts are more optimistic $(e^{KLSS} \gg e^{a})$									
		Market model				Fama–French five-factor model			
Year	Size	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
1985	45	0.007	–0.078	–0.180	–0.500	–0.004	–0.092	–0.200	–0.501
1986	46	–0.049	–0.123	–0.249	–0.640	–0.047	–0.120	–0.254	–0.644
1987	42	–0.043	–0.115	–0.159	–0.450	–0.047	–0.117	–0.255	–0.551
1988	41	–0.046	–0.140	–0.286	–0.588	–0.052	–0.141	–0.327	–0.764
1989	44	–0.034	–0.092	–0.181	–0.449	–0.041	–0.104	–0.198	–0.496
1990	47	–0.041	–0.150	–0.275	–1.113	–0.060	–0.153	–0.302	–1.036
1991	50	–0.046	–0.170	–0.368	–1.349	–0.057	–0.201	–0.479	–1.715
1992	53	–0.027	–0.146	–0.296	–1.099	–0.070	–0.198	–0.450	–1.612
1993	58	–0.055	–0.131	–0.276	–0.562	–0.065	–0.163	–0.354	–1.150
1994	65	–0.067	–0.135	–0.205	–0.504	–0.061	–0.151	–0.348	–1.114
1995	69	–0.053	–0.130	–0.300	–0.636	–0.038	–0.108	–0.310	–0.900
1996	82	–0.082	–0.169	–0.337	–0.860	–0.067	–0.204	–0.504	–1.629
1997	88	0.006	–0.080	–0.273	–0.848	–0.004	–0.088	–0.256	–1.125
1998	95	–0.026	0.001	–0.032	–0.495	–0.025	–0.109	–0.273	–1.099
1999	96	–0.047	–0.102	–0.216	–0.967	–0.015	–0.070	–0.265	–1.168
2000	93	–0.117	–0.311	–0.602	–2.816	–0.133	–0.380	–0.828	–3.537
2001	89	–0.082	–0.046	–0.115	–0.355	–0.041	–0.045	–0.216	–0.497
2002	96	–0.081	–0.185	–0.451	–1.331	–0.073	–0.208	–0.539	–1.840
2003	105	–0.006	–0.078	–0.168	–0.388	0.047	0.031	0.022	–0.114
2004	109	–0.056	–0.213	–0.516	–1.317	–0.044	–0.141	–0.423	–1.089
2005	107	–0.001	–0.027	–0.039	–0.340	–0.003	0.001	–0.047	–0.202
2006	109	–0.046	–0.197	–0.463	–1.175	–0.036	–0.163	–0.413	–1.096
2007	111	–0.022	–0.077	–0.168	–0.303	–0.023	–0.092	–0.203	–0.381
2008	113	–0.013	–0.122	–0.489	–1.530	0.007	–0.067	–0.405	–1.307
2009	120	–0.072	–0.154	–0.414	–0.799	–0.034	–0.091	–0.310	–0.707
2010	127	–0.019	–0.151	–0.449	–1.376	–0.002	–0.103	–0.340	–1.121
2011	121	–0.007	–0.057	–0.154	–0.506	–0.008	–0.056	–0.176	–0.504
2012	121	–0.008	–0.094	–0.283	–0.720	–0.028	–0.100	–0.245	–0.633
2013	125	–0.038	–0.114	–0.319	–0.836	–0.039	–0.128	–0.299	–0.758
2014	127	–0.057	–0.171	–0.311	–0.404	–0.058	–0.191	–0.426	–0.584
Average		–0.041	–0.125	–0.286	–0.842	–0.037	–0.125	–0.321	–0.996
t-stat		*(–7.65)*	*(–11.23)*	*(–11.29)*	*(–8.87)*	*(–6.27)*	*(–9.15)*	*(–11.20)*	*(–8.38)*
Positive		7%	3%	0%	0%	7%	7%	3%	0%
Negative		93%	97%	100%	100%	93%	93%	97%	100%